PROF. KLAUS HENLE (Orcid ID : 0000-0002-6647-5362)
DR. BERND GRUBER (Orcid ID : 0000-0003-0078-8179)

Article type : Research Article

Handling editor: Dr. Anne Chao

Performance of multistate mark-recapture models for temporary emigration in the presence of survival costs

Klaus Henle[1], Bernd Gruber[1,2]

1: *UFZ – Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstr. 15, D-04318 Leipzig, Germany*

2: *Institute for Applied Ecology, Faculty of Applied Science, University of Canberra, Canberra, ACT 2601 Australia*

Corresponding author:

Klaus Henle; e-mail: klaus.henle@ufz.de

Phone: +49-341-2351270

Fax: +49-341-2351470

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/2041-210X.12891

Author Contributions: KH conceived the study; BG wrote the computer program and analysed the models; KH and BG together wrote the manuscript.

Abstract

1. Temporary emigration is widespread in animals and plants and has important implications for ecological processes, evolution, and the conservation of species. It is increasingly studied with capture-mark-recapture (CMR) models. Temporary emigration provides particular challenges to CMR analyses if it involves movement to an unobservable state. A multistate model in which individuals may move between an “observable” and an “unobservable” state (called TE model) was developed for such cases. The model assumes equal survival probability in both states. This assumption may be violated, especially if temporary emigration involves trade-offs between survival and reproduction. A comprehensive assessment of the effects of unequal survival probability on power to detect temporary emigration and on bias and precision of estimates is still needed to understand the applicability and limits of the model.
2. We assessed power to detect temporary emigration for four goodness-of-fit tests and evaluated bias and precision of estimates for the TE model and for its combination with a robust design.
3. Our simulation study, based on 16,650 parameter combinations, shows that temporary emigration is more challenging than currently usually acknowledged. The Tests 2.CT and 2.C are largely independent of the difference in survival probability between the states. In contrast, Test 3.SR is sensitive to the difference in survival probability but

also to emigration probability. Tests 2.C and 2.CT have high power if a large part of the population temporarily emigrates and a large fraction of the individuals return on the next capture occasion. Under this condition, bias is low and precision adequate even if the assumption of equal survival probability is violated. Bias and precision are also satisfactory if the assumption is met but unsatisfactory or unreliable for the remaining parameter space.

4. We conclude that the uncertainties whether an appropriate model was selected and whether the estimates from the selected model may be biased should be clearly communicated and that every endeavour should be made to make the unobservable state observable.

Key words: Bias . Power . Simulation . Survival . Temporary emigration.

Introduction

Temporary emigration, i.e. temporary absence of some individuals from a sampling site, is widespread among animals. In many salamander species, for example, only a subset of the individuals is active and accessible at any time above ground (Bailey *et al.* 2004a). In birds, temporary emigration may occur because of incomplete philopatry to breeding or wintering sites (Hestbeck *et al.* 1991). In several vertebrate groups, individuals may temporarily emigrate to areas where foraging conditions or breeding success are better (e.g. Cooch *et al.* 1999). Likewise, individuals may be temporarily unavailable for capture because they are dormant, in torpor, or hibernate (Kendall *et al.* 1997). Mobile species, such as whales, may temporarily leave and re-enter the study area, as it is often impracticable to carry out the study across the entire range of the population (Whitehead 1990).

Temporary emigration can be of considerable ecological, evolutionary, and conservation biological interest. For example, individuals that skip a breeding opportunity, a phenomenon found in many taxa (Rivalan *et al.* 2005), may be absent from a breeding site (Henle 2005) or may stay in the study area but join a pool of floater that is difficult to sample (Manly *et al.* 1999). Such populations may appear to be stable when in fact they are declining, which may spark intensive controversy in conservation biology as has been the case for spotted owls (*Strix occidentalis*) (Lande 1988) and cormorants (*Phalacrocorax carbo*) (Jepsen & Olesen 2013).

Temporary emigration may bias estimates of survival probabilities from mark-recapture data when using the classical Cormack-Jolly-Seber (CJS) model, which does not account for temporary emigration (Kendall *et al.* 1997). Magnitude and direction of bias depend on the kind of temporary emigration. If all individuals have the same probability of being absent at a given occasion – which is unlikely for many biological situations, but see Pilastro et al. (2003) – temporary emigration is random, and estimates of survival are unbiased (Kendall *et al.* 1997). In contrast, if the probability of being temporarily absent depends on whether an individual was present during the previous occasion, temporary emigration is Markovian ("non-random") and estimates of survival can be substantially biased and even qualitatively different (Kendall *et al.* 1997; Henle 2005).

Advances in mark-recapture models that account for temporary emigration (e.g., Kendall *et al.* 1997; Fujiwara & Caswell 2002; Kendall & Nichols 2002; Bailey *et al.* 2004b) have sparked interest in their application (e.g., Schmidt *et al.* 2002; Bailey *et al.* 2004a). For systems in which temporary emigration involves movement to an unobservable state Fujiwara & Caswell (2002) and Kendall & Nichols (2002) proposed a multistate model termed temporary emigration (TE) model by Schaub *et al.* (2004). The model assumes equal survival probability independent of the state. This assumption is needed to allow parameter identifiability (Schaub *et al.* 2004) and is common to other current models of temporary emigration.

The assumption of equal survival probability is violated in many populations, especially if temporary emigration involves reproduction or takes place in human dominated landscapes. Trade-offs between skipping reproduction and survival is a key issue in life history theories (Roff 1992) though apparent exceptions exist (Reznick *et al.* 2000). In amphibians with a prolonged breeding season, males may temporarily leave the breeding site to replenish resources, thus temporarily missing breeding opportunities (von Lindeiner 1992). Likewise, juvenile lizards may be pushed to suboptimal habitats by adults and thus may experience increased predation pressure (Brandl & Völkl 1988).

Survival costs may be substantial. For example, in a dry year, male tiger salamanders (*Ambystoma tigrinum*) that bred had a survival probability of only approx. 0.35 whereas it was approx. 0.95 in those that skipped breeding (Church *et al.* 2007). Large differences also exist in various temperate snakes, which may approach semelparity (Bonnet *et al.* 2002); in semelparous species all reproducing individuals die whereas non-reproducing individuals may have a high survival probability (e.g. Cogger 1974).

Movement itself can be costly (Gruber & Henle 2008). In human dominated landscapes mortality while crossing roads may reach 70-100% in amphibians, turtles, snakes, and even in fast moving species, such as otters (*Lutra lutra*) (Frank *et al.* 2002; Andrews *et al.* 2008). Finally, in fire prone systems individuals that are active may be killed or permanently dispersed while dormant ones or those that temporarily emigrate to inaccessible refuges within or outside the affected sites may survive (Driscoll & Roberts 1997).

Because survival frequently differs in the observable and unobservable state, it is of considerable interest to know the power of statistical methods to detect temporary emigration and how well mark-recapture models that include temporary emigration perform when the assumption of equal survival probability is not met. No specific test for temporary emigration has been developed and most of the suggested tests are based on different partitioning of the overall goodness-of-fit (GOF) test of the classical CJS model (Cooch *et al.* 1999; Manly *et al.*

1999; Schmidt *et al.* 2002). This procedure consists of two overall tests each with two test components. Test 2 tests for differences in catchability or survival among release cohorts, whereas Test 3 tests for differences in catchability or survival for individuals with different capture histories within a release cohort. A full interpretation of these tests in the context of temporary emigration is provided in Supplement S7.

Authors disagree regarding which component of the overall GOF developed by Burnham *et al.* (1987) should be used (Pradel *et al.* 1997; Cooch *et al.* 1999; Manly *et al.* 1999; Schaub *et al.* 2004). Conclusions about power range from very high (Schaub *et al.* 2004) to very low (Manly *et al.* 1999). Unfortunately, all authors used different parameter combinations and assessed only a very small part of the entire parameter space, making it impossible to reconcile the diverging conclusions and to decide which method to use and how reliable it will be.

In terms of bias, Schaub *et al.* (2004) were the first to systematically analyse it for the TE model. This model is suitable for the most common sampling scheme, in which individuals are sampled once at each primary occasion between which the population is open to losses and gains. In the TE model, individuals may move between an "observable" and an "unobservable" state. Individuals in a sampling area are in the observable state whereas moving out of the sampling site is equivalent to becoming a temporary emigrant. In principal it is possible to allow time dependence in the TE model, but for almost all combinations of time dependency the movement parameters are not estimable (Kendall & Nichols 2002). Schaub *et al.* (2004) restricted their assessment to scenarios in which individuals that are present and those that are temporarily absent have the same survival probabilities and concluded that the model performs well under first order Markovian temporary emigration when survival and recapture probabilities are high.

Our study was motivated by the divergent opinions about the suitability of different GOF to test for temporary emigration (see above) and by disputes about the reliability of parameter estimates for a breeding population of the common toad (*Bufo bufo*) (Henle 2001, 2005; Schaub et al. 2004). Here we address three interlinked issues. First, we provide a comprehensive assessment of the power of different components of the overall goodness-of-fit test to the CJS model to detect temporary emigration when the assumption of equal survival is violated. This is critical to assess the risk of falsely assuming absence of temporary emigration and for the decision whether a standard CJS model or a TE model should be used. In particular, we include the tests that were used or recommended by Manly *et al.* (1999) or Schaub *et al.* (2004) that lead these authors to opposing conclusions. Second, we analyse bias and precision of parameter estimates for the TE model under survival costs in the observable state. Third, we repeated the analyses with a robust design, in which individuals are sampled additionally during secondary periods within primary periods, which is superior to models without secondary periods, because it allows an independent estimate of capture probabilities (Kendall 2004). Finally, we provide recommendations to decide which test to use to be able to detect temporary emigration, identify the range of parameter values for which bias may be small, and make recommendations how to plan studies that may overcome problems created by an unobservable state.

Simulation methods

TE Model

The TE model is a two-state transition model with one state being unobservable. Individuals move with transition probabilities ψ_{OU} from the observable state to the unobservable state and ψ_{UO} from the unobservable state to the observable state. We were interested in the performance of the model to estimate its parameters when the assumption of equal survival probabilities in the observable and the unobservable state during one time step is not met ($S_U \neq S_O$).

We simulated capture-recapture studies, systematically varying survival in the observable and unobservable state, to identify cases, where temporary emigration can be detected. We ran 16,650 different combinations of parameters to sample the parameter space as comprehensively as possible. We varied S_O from 0.1 to 0.9 in steps of 0.2 and S_U from 0.1 to 1 in steps of 0.1. We omitted cases where $S_U < S_O$, because of space limits and because our interest in the topic derived from controversies about breeding populations, in which presence in the observable state incurs survival costs. This results in a total of 30 combinations regarding S_O and S_U. Readers interested in other scenarios can use our simulation tool to assess them (Supplement S3).

We varied recapture rate from 0.2 to 1 by steps of 0.2 (5 levels) and the transition parameters ψ_{OU} and ψ_{UO} from 0 to 1 by steps of 0.1 (121 combinations). We collapsed the ten combinations where ψ_{OU} was zero and ψ_{UO} was varied from 0 to 1 to one scenario ($\psi_{OU} = 0$), as it does not make sense to vary ψ_{UO} when ψ_{OU} is zero, because no marked animal would leave the capture area. The total number of different parameter combination is thus 30 x 5 x 111 = 16,650 scenarios.

To allow direct comparison with the results of Schaub *et al.* (2004) we assessed the scenarios using their suggested approach called 'analytical-numerical approach', which was also recommended by Burnham *et al.* (1987). In this approach, the expected numbers of individuals are calculated for each capture history with the known parameters, which are then used as 'actual' capture data to estimate parameters and their coefficient of variation. Like Schaub *et al.* (2004) we considered eight capture occasions and 500 newly released individuals at each occasion. For generating capture histories we created a C++ program with a user-friendly interface (Supplement, S3). The program can be used to simulate capture histories under the assumption of the TE model to test power and potential bias of a planned study for expected parameter combinations. We estimated the model parameters with program MARK (White & Burnham 1999).

For the 16,650 scenarios we computed four GOF tests to compare their performance to detect temporal emigration: Test 2.CT from program U-CARE (Choquet *et al.* 2001) and Tests 2.C, 3.SR, and 3.SM from program RELEASE (Burnham *et al.* 1987). Test 2.CT was used by Schaub *et al.* (2004); we included the latter three tests because Manly *et al.* (1999) suggested that these partial tests of Test 2 and Test 3 should be more powerful than the overall tests to detect temporary emigration. We had to implement Test 2.CT by ourselves as no command line tool was available for this test and having 16,650 tests to be done, automation was necessary. In the results, we express power as the percentage of scenarios in which the GOF resulted in a detection of temporary emigration at $\alpha < 0.05$.

We evaluated absolute bias by comparing the estimated parameter $\hat{\theta}$ to the value used in the simulation using $bias(\hat{\theta}) = \hat{\theta} - \theta$ and precision by calculating the coefficient of variation (*CV*) of the estimated parameter $CV(\hat{\theta}) = \frac{\sigma(\hat{\theta})}{\mu(\hat{\theta})}$, where σ and μ are the standard error and the mean of the estimated parameter, respectively. For bias of survival, we determined the actual survival rate across all individuals (S_t^*) by counting all surviving individuals for each time period. We used its average across the seven time periods (S*) to calculate the bias of $\hat{S}$ as ($bias(\hat{S}) = \hat{S} - S^*$).

Robust design

We also assessed bias and CV of parameter estimates when adding closed secondary periods to estimate capture probability (*p*), hence implementing a robust design (Pollock *et al.* 1990).

Using these estimates of p we estimated the remaining parameters of the standard TE model based on the primary periods. If sample sizes are large as in our simulation, the analytical-numerical approach will provide an estimate of p that is very close to its expected value. Therefore, we obtained estimates, their *bias* and *CV* as explained for the standard TE model, fixing p to its true value. Note that variances will be underestimated with this approach.

Results

Parameter estimates, their bias and CV as well as the ability of four goodness-of-fit tests to detect temporary emigration are presented for each of the 16,650 parameter combinations in the online supporting material S5. Bias and power depended on interactions of the model parameters and few simple rules emerged. In the following, we therefore summarize the results in terms of deviations from the assumption of the TE model of equal survival probabilities in the observable and the unobservable state (S_U-S_O) and in terms of the deviations from random temporary emigration [$\psi_{ou}-(1-\psi_{uo})$]. Temporary emigration is random if $\psi_{ou}-(1-\psi_{uo})=0$.

Assumptions of the TE model met

Detection of temporary emigration

Results of the scenarios in which the assumption of the TE-model were met (i.e., S_u=S_o) are presented in Figure S1. Test 2.CT showed its best performance when both ψ_{OU} and ψ_{UO} were high. Test 2.C showed a similar performance pattern but temporary emigration was consistently detected for a higher percentage of scenarios. Test 3.SR achieved significance only when ψ_{OU} was very high (but smaller than 1) and Test 3.SM only for a narrow range of very high values of ψ_{OU} (0.7-0.9). ψ_{UO} had only a marginal effect on the performance of both tests.

Bias and coefficient of variation (CV) of parameter estimates

Bias of parameter estimates is shown in Figures 2-5 and S2.1-S2.3, bottom lines. All parameter estimates were unbiased, independent of the values of other parameters and similar whether a robust design was used or not. An exception was when emigration approached permanency, for which no reliable estimates were possible. A further exception was when there was no temporary emigration for which at least one parameter was strongly biased for most parameter combinations, with survival estimates tending to be less affected than estimates of migration probabilities (Supplement S5). In addition, without a robust design, reliable parameter estimation was not possible when temporary emigration was random because of parameter confounding.

The CV for survival probability, emigration probability, and return probability was low, if $\psi_{ou} - (1 - \psi_{uo})$ was higher than 0, 0.4, and 0.6, respectively (-0.3, 0, and 0.2, when using a robust design).

Assumptions of the TE model not met

Detection of temporary emigration

Tests 2CT and 2.C were highly sensitive to deviations from random temporary emigration and depended only weakly on differences in survival probabilities (Figs. 1a,b). If $\psi_{ou} - (1 - \psi_{uo}) > 0.3$, most parameter combinations achieved significance but the closer emigration was to random, the fewer parameter combinations achieved it. For the remaining parameter space only few parameter combinations achieved significance and only when survival probability was much higher in the unobservable state than in the observable state.

Test 3.SR was sensitive to the difference in survival probability and depended only weakly on the degree of deviation from random temporary emigration (Fig. 1c). If $S_u - S_o > 0.6$, most parameter combinations resulted in significant tests. However, when emigration approached permanency; power was low. Under moderate differences in survival probability (< 0.2) the test was not significant for the majority of parameter combinations. Test 3.SM had high power in a narrow range of the parameter space, in which the difference between survival probabilities and the deviation from random temporary emigration were high, except when emigration approached permanency (Fig. 1d). See Table S1 for the percentage of scenarios for which the GOF tests were significant for different values of ψ_{OU} and ψ_{UO}.

Bias and coefficient of variation (CV) of parameter estimates

Standard TE model

Bias and CV of estimates of survival across all individuals ($\hat{S}^*$). Bias of $\hat{S}^*$ (overestimation) increased with the difference in survival probability (Fig. 2a; Supplement S4: $r = 0.5$) and decreased with survival probability in the observable state (Supplement S4: $r = -0.43$). When emigration probability was higher than the probability to stay emigrated, bias was low and positive, except for cases in which temporary emigration was random (hatched central part of Fig. 2a). Under random temporary emigration estimates were erratic, i.e., parameter combinations for which parameter estimates were highly biased were very close to parameter combinations, for which there was no bias (Supplement S6). When emigration probability was smaller than the probability to stay emigrated, low bias was achieved only when the difference in survival probability was limited (< 0.3).

Under permanent emigration $\hat{S}^*$ was unbiased, severely overestimated or underestimated, depending on emigration probability and the difference in survival probability.

The CV of $\hat{S}^*$ was low, if the emigration probability was higher than the probability to stay emigrated. It was always high when the probability to stay emigrated was higher than emigration probability (Figs. 3b).

In summary, one may get relatively precise and unbiased estimates if the emigration probability was higher than the probability to stay emigrated.

Bias and CV of estimates of the probability to emigrate ($\hat{\psi}_{ou}$). Bias of (overestimation) increased moderately with the difference in survival probability (Fig. 3a, Supplement S4: *r* = 0.35) and decreased moderately with ψ_{OU} (Supplement S4: *r* = -0.38). When emigration probability was higher than the probability to stay emigrated, bias was low and positive. When temporary emigration was random no reliable estimates were possible (hatched central part of Fig. 3a). When emigration probability was smaller than the probability to stay emigrated, bias was low only when the difference in survival probability was limited (< 0.2).

Under permanent emigration $\hat{\psi}_{ou}$ was unbiased, severely overestimated or underestimated, depending on emigration probability and the difference in survival probability.

CV of $\hat{\psi}_{ou}$ showed a complex pattern and was low when the probability to emigrate was higher (approx. ≥ 0.4) than the probability to stay emigrated or when the former was much lower than the latter and at the same time the difference in survival probability was at least 0.4 (Fig. 3b). For all other parameter combinations CV was very high.

In summary, reliable and precise estimates were possible when $\psi_{ou} - (1 - \psi_{uo}) \geq 0.3$.

Bias and CV of estimates of the probability to return from emigration ($\hat{\psi}_{uo}$). Bias of $\hat{\psi}_{uo}$ (underestimation) decreased moderately with ψ_{UO} (r = -0.38) and bias in $\hat{S}^*$ (r = -0.35) and considerably (r = -0.47) with bias in $\hat{\psi}_{ou}$ (Supplement S4). Bias was high when the difference in survival probability was high (≥ 0.6-0.7) and at the same time emigration probability was similar to the probability to stay emigrated (Figs. 4a). In addition, estimates were unreliable if temporary emigration was random. Under permanent emigration $\hat{\psi}_{uo}$ was either unbiased or severely overestimated, depending on the values of the other parameters.

CV of $\hat{\psi}_{uo}$ was very high unless emigration probability was very high and the probability to stay emigrated was very low (difference > 0.7) (Figs. 4b). In summary, reliable and robust estimates were obtained only when the difference between emigration probability and the probability to stay emigrated was high (>0.5) and when the difference in survival probability was limited (< 0.5).

Bias and CV of estimates of capture probability ($\hat{p}$). Bias in $\hat{p}$ was independent of true survival probabilities and their difference, capture probability, migration probabilities and their difference (Supplement S4). Estimates of capture probability were practically unbiased for about 90% of the parameter space explored (Fig. 5a, Supplement 5). The remaining 10% were scenarios in which reliable estimates were not possible because of parameter confounding.

The CV of $\hat{p}$ was high unless the difference in the transition probabilities was very high (>0.7; Fig. 5b).

Interrelationships of biases in parameter estimates. Biases of parameter estimates were correlated with the bias of estimates of another parameter in four cases (three times strongly and once moderately; Supplement S4). Bias in the estimates of $\hat{S}^*$ was strongly correlated with bias in $\hat{\psi}_{ou}$ (r = 0.56) and moderately negatively with bias in $\hat{\psi}_{uo}$ (r = -0.35). Bias in

$\hat{\psi}_{ou}$ and $\hat{\psi}_{uo}$ showed a moderately strong negative relationship (r = -0.47) and $\hat{\psi}_{ou}$ was further positively correlated with bias in $\hat{p}$ ($r = 0.44$).

Robust design

Bias for all parameter estimates was very similar to those obtained with the TE model without a robust design. The only main difference was that the robust design allowed reliable parameter estimates, when temporary emigration was random (Figs. S2.1a-S2.3a), which was not possible without a robust design (Figs 2a-5a). The pattern of CV was also similar to the TE model without a robust design for $\hat{S}^*$ and for $\hat{\psi}_{uo}$ (Figs. S2.1b, S2.3b) but the values were slightly smaller. However, it differed substantially for emigration probability (Fig. 3b, S2.2b). While it was high across a large part of the parameter space in the TE model without a robust design, it was low across most of the parameter space examined when a robust design was used. In the latter case it was high only when the difference in survival probability was low (< 0.3) and concomitantly the probability to emigrate was lower than the probability to stay emigrated (Fig. S2.2b).

Discussion

Power to detect temporary emigration

Most authors suggested using the full testing procedure for CJS-models, or parts thereof, developed by Burnham *et al.* (1987) to test for temporary emigration. Opinions about the power of these tests to detect temporary emigration differ among authors. Simulating a recapture study with higher order Markovian emigration in the spotted owl (*Strix occidentalis*), Manly *et al.* (1999) noticed that the overall Tests 2 and 3 had low power and required large amounts of recapture data to detect temporary emigration. They and Cooch *et* *al.* (1999) suggested that partial tests should be more powerful than the overall tests. Cooch *et al.* (1999) further regarded Test 2.C as diagnostic for temporary emigration. However, the test can be significant even when there is no temporary emigration (in 31% of the parameter combinations assessed by us; Table S1b). Using partial Test 2.CT, Schaub et al. (2004) obtained high power in their simulation of a virtual species under first-order Markovian temporary emigration and identical survival probability in the observable and the unobservable state.

We compared for the first time the power of these GOF tests for the same data sets. Our simulations, based on a far larger range of parameter combinations than previous simulations, showed that Tests 2.C and 2.CT depend on the difference between emigration probability and the probability to stay emigrated and are largely independent of the difference in survival probability. They perform well, if the probability to emigrate is at least 0.2 higher than the probability to remain emigrated. Notwithstanding, Test 2.C detected temporary emigration in at least 80% of the scenarios only if both ψ_{OU} and ψ_{UO} were high, i.e., if at least 50% of the individuals emigrated temporarily and many of the emigrants returned at the next occasion. Else, significant results can be expected only for a limited parameter space, when the difference in survival is very high. In this case, Test 2.C performs better than Test 2.CT. However, test 2.CT has the big advantage that it was never significant if there was no temporary emigration.

In contrast to Tests 2.C and 2.CT, Test 3.SR is primarily sensitive to the difference in survival probability and thus might be useful to test for the assumption of equal probability of survival in the observable and unobservable state. Survival difference must be large (>0.3) for Test 3.SR to detect it in more than 50% of the scenarios. Test 3.SM is significant for fewer parameter combinations across the whole parameter space examined than Test 3.SR. Therefore, it is less suitable than Test 3.SR in the context of temporary emigration. Both tests are also sensitive to the transition probabilities and may be significant even when survival probabilities do not differ (Fig. S1).

Under random temporary emigration Test 2.CT was never significant. This is in line with the results of Schaub *et al.* (2004). The other three tests were significant for a subset of the scenarios of random emigration, with 3.SR being significant in > 50% of the scenarios. The results were the same for all scenarios of permanent emigration.

Our results confirm the conclusion of Manly *et al.* (1999) that the components of Test 3 have limited power to detect temporary emigration. However, contrary to Schaub *et al.* (2004), we obtained high power for Test 2.CT only for a limited range of the parameter space (at least 0.2 higher than the probability to stay emigrated). The difference between our results and that of Schaub *et al.* (2004) can be explained by one additional level of realism, i.e. complexity, of our model: the incorporation of differential survival in the observable and unobservable state. While being more general, our model still is rather restrictive. It neither includes temporal variability in any parameter, nor individual heterogeneity, or behavioural response to capture. It also does not incorporate higher order Markovian processes. The inclusion of such additional realism will likely increase the difficulty to detect temporary emigration (Manly *et al.* (1999; Henle 2005) and to separate it from other processes that may create heterogeneity in survival and capture probability.

In conclusion, one can have confidence in good power of GOF tests for temporary emigration, if the probability of temporary emigration is at least 30% and much higher than the probability to stay emigrated (i.e., most individuals return at the next capture occasion). Additionally, one must distinguish between Test 2.C/Test 2.CT, which are primarily transition tests, and Test 3.SR, which primarily tests for the assumption of equal survival probability – but take care that the latter is also sensitive to emigration probability.

Bias in parameter estimates

If GOF tests did not detect temporary emigration and CJS models without temporary emigration are used, estimates can be biased. For example, in the case of random temporary emigration capture probability and emigration probability are confounded (Kendall *et al.* 1997). It still remains to be systematically explored how strong bias is when survival probabilities differ among states and for different combinations of emigration and re-immigration probabilities.

Our results show that the estimates are also unreliable for the standard TE model when temporary emigration is random. Bias can also be substantial in case of higher order Markovian temporary emigration as shown for one specific case by Henle (2005) but still needs to be explored across a wider parameter space.

If temporary emigration is detected or assumed and the TE model is used for parameter estimation, bias in all parameters is always very small if there is a low difference in survival probability (S_U-$S_O \leq 0.1$), thus confirming Schaub *et al.* (2004). For this parameter space, CV is also low for all parameter estimates, as long as the probability of temporary emigration is higher than the probability to stay emigrated [$\psi_{ou} - (1 - \psi_{uo}) > 0.2$].

Bias of $\hat{S}^*$ scales roughly with S_U-S_O and becomes considerable if S_U-$S_O > 0.2$ unless the probability to emigrate is at least 0.2 higher than the probability to remain emigrated. Under the latter condition, CV($\hat{S}^*$) and bias of the remaining parameter estimates are also limited, but CVs of the remaining parameter estimates, especially of $\hat{\psi}_{uo}$, are small only under larger differences in transition probabilities. A large CV of $\hat{\psi}_{uo}$ is to be expected given that this is the most indirectly estimated parameter (Kendall *et al.* 1997). For the remaining parameter space, estimates and/or CVs tend to be strongly biased and/or very wide, respectively.

When temporary emigration is random, emigration probability was confounded with capture probability as is the case when instead of the TE model a CJS model without temporary emigration is used for parameter estimation (Kendall & Nichols 1995). For permanent emigration estimates were highly unreliable as minor changes in a single parameter value changed the results from being unbiased to being severely biased and vice versa in an unpredictable direction.

When the standard TE model was combined with a robust sampling design results were similar in terms of bias in about 90% of the cases but different in terms of precision. This is expected as fewer parameters have to be estimated when independent information is available to estimate capture probability. As an additional advantage, the robust design extended the parameter space in which both relatively unbiased and precise estimates were possible to random temporary emigration and to cases in which the probability to stay emigrated it slightly higher (≤ 0.2) than the probability to temporarily emigrate.

We caution that our graphs are averages across many parameter combinations and bias for some combinations may still be substantial. Therefore, the table of all estimates (Supplement S5) should be always consulted as well. A further caveat is that currently there is no statistical diagnostic to differentiate whether a test is significant because emigration probability is much higher than the probability to stay emigrated or because of the opposite combined with large differences in survival probability. The bad news is that under the latter condition estimates are either substantially biased or/and have very large CVs. Survival (across all individuals) is substantially overestimated, which is particularly problematic in the management of rare or endangered species or when planning for sustainably harvesting a population (Jepsen & Olesen 2013). Harvesting incurs mortality costs in the observable state and overestimation of survival increases with the intensity of harvest.

Conclusions

Our simulation study shows that temporary emigration is more challenging to the analysis of populations than currently usually acknowledged but the good news is that if the limitations are carefully considered and there is good biological knowledge that all assumptions apart from equal survival probability in the observable and the unobservable state are met, one may get reliable results. Notwithstanding, the conclusion of Schaub *et al.* (2004) that power is usually high holds only for a limited part of the parameter space: (1) a large part of the population must temporarily emigrate and most individuals return on the next capture occasion or (2) survival probability in the unobservable state is much higher than in the observable state and the probability to stay emigrated is moderately higher than the probability to emigrate (Kendall & Nichols 2002; this study). In addition, temporary emigration must be first-order Markovian and a large number of individuals must be marked (Manly *et al.* 1999; Kendall & Nichols 2002). First order Markovian temporary emigration is unlikely when studying territorial species with floater systems (Manly *et al.* 1999) but may hold for other systems. The good news is that if condition (1) holds, the likelihood to get a significant result from Test 2.C is high and parameter estimates have low bias (if $\psi_{ou} - (1 - \psi_{uo}) > 0.2$) and are rather precise (if $\psi_{ou} - (1 - \psi_{uo}) > 0.4$), even if the assumption of equal survival probability in the observable and the unobservable state is not met.

Falsely using a CJS model without temporary emigration because the GOF-tests were not significant may also lead to major bias in parameters estimates (Kendall *et al.* 1997; Henle 2005). Unfortunately, estimates of the TE model are also unreliable if there is no temporary emigration. Test 2.CT is the only test that is never significant in this case and therefore should always be included in testing. As Tests 2.CT and 2.C depend primarily on differences in transition probabilities and Test 3.SR primarily on differences in survival probabilities, we recommend combining one of the former two and the latter tests.

The uncertainties whether an appropriate model was selected and whether the estimates from the selected model may be biased or not are uncertainties that are not covered by just reporting confidence intervals. Therefore, they should be clearly communicated (Henle 2005).

Likewise, one should have good biological arguments that temporary emigration is not higher order Markovian and that the assumption of parameter constancy across years is either reasonable or the estimate is a good approximation to the geometric mean of the estimated parameters across the sampling years. Violation of additional assumptions, such as time constancy, likely will increase bias. Likewise, violation of the assumption of homogeneous capture probability may further increase bias, though in the standard Jolly-Seber model it produces only small bias in survival estimates (Williams *et al.* 2002). CMR analyses have seen tremendous advances in recent decades but we recommend that the interested community now moves into a phase of similar advances in comprehensive assessments of the robustness and reliability of parameter estimates under the full range of biologically plausible deviations from model assumptions.

While our tables and figures provide a first indication of power and bias that may be achieved in a field study with large capture data depending on expected parameter values, we developed a computer programme (for Windows) that allows simulating and analysing capture histories (using MARK) under the TE model for further parameter combinations, e.g. if survival is lower in the unobservable state, if fewer or more capture occasions are planned, and for smaller populations. The program can be downloaded from the authors' webpage (Supplement S3) and used to test the power and potential bias and precision in parameter estimates of a planned field study.

We agree with Schaub *et al.* (2004) that the onus is on a carefully designed sampling scheme that includes a robust design but this is not sufficient. While the robust design does not reduce bias, it increases the precision of parameter estimates and allows reliable and precise estimates of survival probability – if the difference in survival probabilities is limited – and of emigration probability under random temporary emigration. As the parameter space is limited in which GOF tests are likely to detect temporary emigration and parameter estimates are unbiased and precise, we recommend that any effort is undertaken to make the unobservable state observable, so that multistate models without an unobservable state (e.g., Nichols *et al.* 1994) can be used. This may be achieved by surveying a strip around the study site to sample temporarily emigrated individuals (e.g., Gruber & Henle 2008) or by combining the mark-recapture study with a telemetry study to allow following temporarily emigrated individuals and quantify mortality (Kendall 2004). One should also attempt to collect data that allow an independent estimate of mortality while in the observable state, e.g., by intercepting migrating amphibians with fences when they enter and leave breeding pond and by determining capture probability at the fences. Recapture models exist that allow parameter estimation for such cases even when survival probability differs between the observable and unobservable state (Bailey *et al.* 2004b) – as long as there is no mortality while moving to the unobservable state. While such approaches may require considerable effort and may be technically challenging, only then will our understanding of temporary emigration processes and their importance in ecology, evolution, and species conservation make major advances.

Acknowledgements We thank Marlies Uhlig (Leipzig) for support in collating the references. Annegret Grimm (Leipzig), Jean-Baptiste Mihoub (Leipzig), Michael Schaub (Sempach) and two anonymous reviewers provided valuable comments to previous versions of the manuscript that helped substantially to improve its clarity.

Data Accessibility. A comprehensive table that entails detailed results for all 16550 parameter combinations and a simulation program that can be used to explore user defined parameter combinations, can be found on Github: https://github.com/green-striped-gecko/SimulaTEr, https://doi.org/10.5281/zenodo.854832

References

Andrews, K.M., Gibbons, J.W. & Jochimsen, D.M. (2008) Ecological effects of roads on amphibians and reptiles: a literature review. *Urban herpetology* (eds J.C. Mitchell, R.E. Jung Brown & B. Bartholomew), pp. 121-143. Society for the Study ofAmphibians and Reptiles, Salt Lake City.

Bailey, L.L., Simons, T.R. & Pollock, K.H. (2004a) Spatial and temporal variation in detection probability of *Plethodon* salamanders using the robust capture-recapture design. *Journal of Wildlife Management,* 68, 14-24.

Bailey, L.L., Kendall, W.L., Church, D.R. & Wilbur, H.M. (2004b) Estimating survival and breeding probability for pond-breeding amphibians: a modified robust design. *Ecology,* 85, 2456-2466.

Brandl, R. & Völkl, W. (1988) Tail break rate in the Madeiran lizard (*Podarcis dugesii*). *Amphibia-Reptilia*, 9, 213-218.

Bonnet, X., Lourdais, O., Shine, R. & Naulleau, G. (2002) Reproduction in a typical capital breeder: costs, currencies, and complications in the aspis viper. *Ecology*, 83, 2124-2135.

Burnham, K.P., Anderson, D.R., White, G.C., *et al.* (1987) Design and analysis methods for fish survival experiments based on release-recapture. *American Fisheries Society Monograph*, 5, 1-437.

Choquet, R., Reboulet, A.M., Pradel, R. & Lebreton, J.D. (2001) *U-CARE (Utilities–capture–recapture) user's guide. Typescript*. CEFE/CNRS, Montpellier, France. http://www.cefe.cnrs.fr/fr/recherche/bc/bbp/1045-desc/264-logiciels [accessed November 2016]

Church, D.R., Bailey, L.L., Wilbur, H.M., *et al.* (2007) Iteroparity in the variable environment of the salamander *Ambystoma tigrinum*. *Ecology*, 88, 891-903.

Cogger, H.G. (1974) Thermal relations of the mallee dragon *Aphibolurus fordi* (Lacertilia: Agamidae). *Australian Journal of Zoology*, 22, 319-339.

Cooch, E.G., Lank, D.B. & Rockwell, R.F. (1999) Body size and age of recruitment in snow geese *Anser c. caerulescens*. *Bird Study*, 46S, S112-S119.

Driscoll, D.A. & Roberts, D. (1997) Impact of fuel reduction burning on the frog *Geocrinia lutea* in southwest Western Autralia. *Austral Ecology* 22, 334-339.

Frank, K., Eulberg, P., Hertweck, K. & Henle, K. (2002) A simulation model for assessing otter mortality due to traffic. *IUCN Otter Specialist Group Bulletin*, 19A, 64-68.

Fujiwara, M. & Caswell, H. (2002) A general approach to temporary emigration in mark–recapture analysis. *Ecology*, 83, 3266–3275.

Gruber, B. & Henle, K. (2008) Analysing the effect of movement on local survival: a new method with an application to a spatially structured population of the arboreal gecko *Gehyra variegata*. *Oecologia*, 154, 679-690.

Henle, K. (2001) Pit-falls in data analysis of amphibian breeding cycles and mortality patterns: a comment on Schmidt and Anholt (1999). *Amphibia-Reptilia*, 22, 379-386.

Henle, K. (2005) Analysis of recapture data from breeding populations of amphibians: On temporary emigration, model assumptions, bias, and common toads. *Amphibia-Reptilia*, 26, 7-16.

Hestbeck, J.B., Nichols, J.D. & Malecki, R.A. (1991) Estimates of movement and site fidelity using mark-resight data of wintering Canada geese. *Ecology*, 72, 523-533.

Jepsen, N. & Olesen, T. (2013) Cormorants in Denmark – re-inforced management and scientific evidence. *Human-wildlife conflicts in Europe — Fisheries and fish-eating vertebrates as a model case* (eds R. Klenke, I. Ring, A. Kranz, *et al.*). Springer, Heidelberg.

Kendall, W.L. (2004) Coping with unobservable and mis-classified states in capture-recapture studies. *Animal Biodiversity and Conservation*, 27.1, 97-107.

Kendall, W.L. & Nichols, J.D. (2002) Estimating state-transition probabilities for unobservable states using capture–recapture/resighting data. *Ecology*, 83, 3276–3284.

Kendall, W.L., Nichols, J.D. & Hines, J.E. (1997) Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology*, 78, 563-578.

Lande, R. (1988) Demographic models of the northern spotted owl (*Strix occidentalis caurina*) *Oecologia*, 75, 601-607.

Manly, B.F.J., McDonald, L.L. & McDonald, T.L. (1999) The robustness of mark-recapture methods: A case study for the northern spotted owl. *Journal of Agricultural Biological and Environmental Statistics*, 4, 78-101.

Pilastro, A., Tavecchia, G. & Marin, G. (2003) Long living and reproduction skipping in the fat dormouse. *Ecology*, 84, 1784-1792.

Pollock, K.H., Nichols, J.D., Brownie, C. & Hines, J.E. (1990) Statistical inference for capture-recapture experiments. *Wildlilfe Monographs*, 107, 1-97.

Pradel, R., Hines, J.E., Lebreton, J.D. & Nichols, J.D. (1997) Capture-recapture survival models taking account of transients. *Biometrics*, 53, 60-72.

Reznick, D., Nunney, L. & Tessier, A. (2000) Big houses, big cars, superfleas and the costs of reproduction. *TREE*, 15, 421-425.

Rivalan, P., Prévot-Julliard, A.-C., Choquet, R., *et al.* (2005) Trade-off between current reproductive effort and delay to next reproduction in the leatherback sea turtle. *Oecologia*, 145, 564-574.

Roff, D.A. (1992) *The evolution of life histories. Theory and analysis.* Chapman & Hall, London.

Schaub, M., Gimenez, O., Schmidt, B.R. & Pradel, R. (2004) Estimating survival and temporary emigration in the multistate capture–recapture framework. *Ecology*, 85, 2107-2113.

Schmidt, B.R., Schaub, M. & Anholt, B.R. (2002) Why you should use capture-recapture methods when estimating survival and breeding probabilities: on bias, temporary emigration, overdispersion, and common toads. *Amphibia-Reptilia*, 23, 375-388.

von Lindeiner, A. (1992) Untersuchungen zur Populationsökologie von Berg-, Faden- und Teichmolch (*Triturus alpestris* L., *T. helveticus* Razoumowski, *T. vulgaris* L.) an ausgewählten Gewässern im Naturpark Schönbuch (Tübingen). *Jahrbuch für Feldherpetologie*, 3, 1-117.

Wiegand, K., Henle, K. & Sarre, S. (2002) Extinction and spatial structure in simulation models. *Conservation Biology*, 16, 117-128.

White, G.C. & Burnham, K.P. (1999) Program MARK: survival estimation from populations of marked animals. *Birds Study*, 46, S120-139.

Whitehead, H. (1990) Mark-recapture estimates with emigration and re-immigration. *Biometrics,* 46, 473-479.

Williams, B.K., Nichols, J.D.. & Conroy, M.J. (2002) *Analysis and management of animal populations.* Academic Press, New York.

List of Figures 286 Wörter

probabilities in/from the observable (S_O, ψ_{OU}) and in/from the unobservable state (S_U, ψ_{UO}). Hatched: parameter space for which parameters could not be identified correctly because of parameter confounding.

Fig. 1 Contour plots for the power of four tests for combinations of differences in survival and transition probabilities.

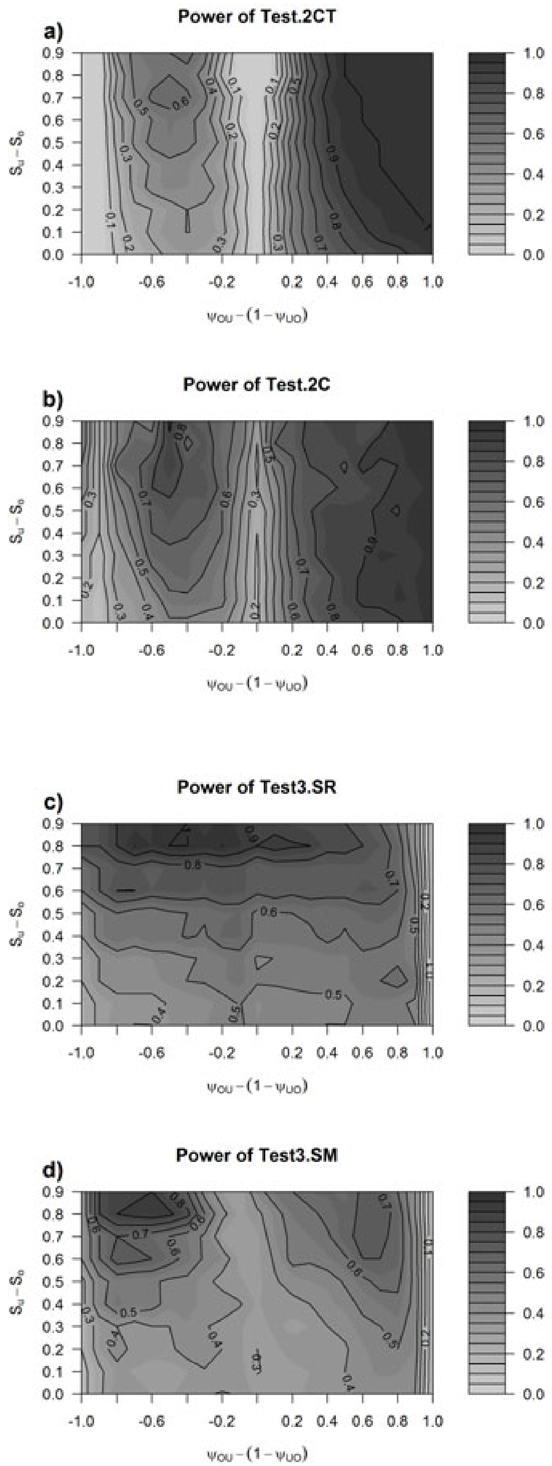
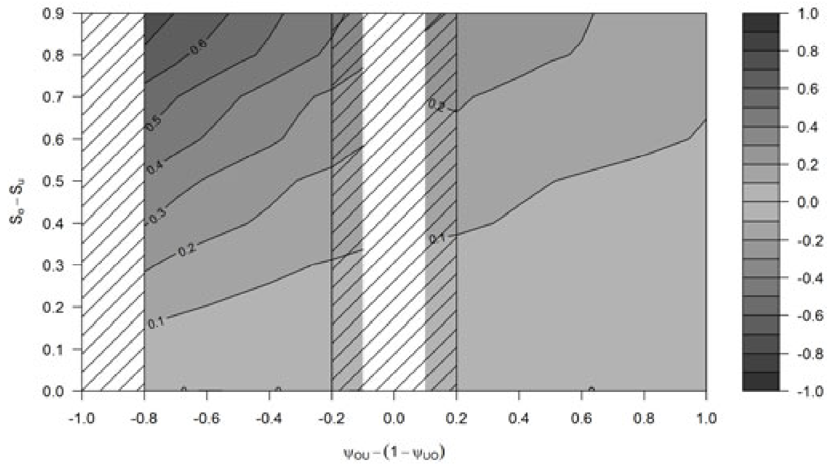


Figure 2 a) Bias and b) CV in S*, TE model.

a)



b)

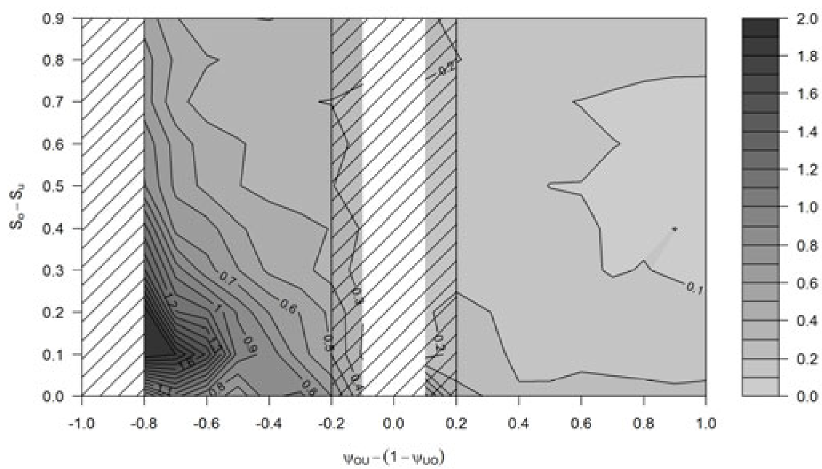
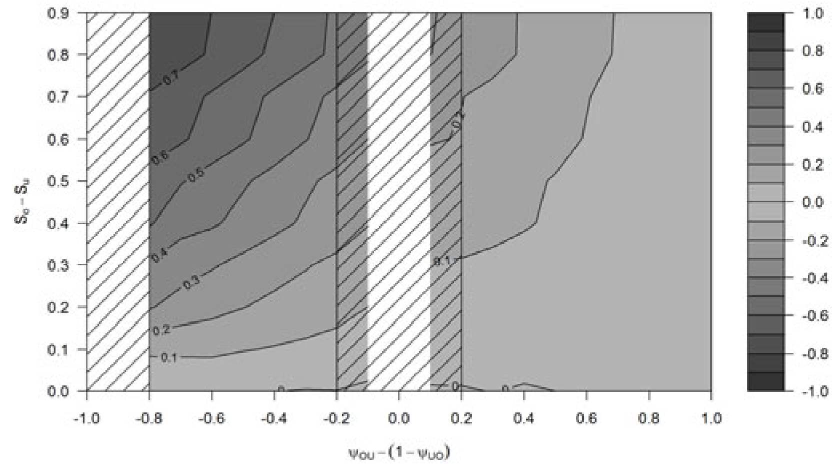


Figure 3 a) Bias and b) CV in ψ_{OU}, TE model

a)



b)

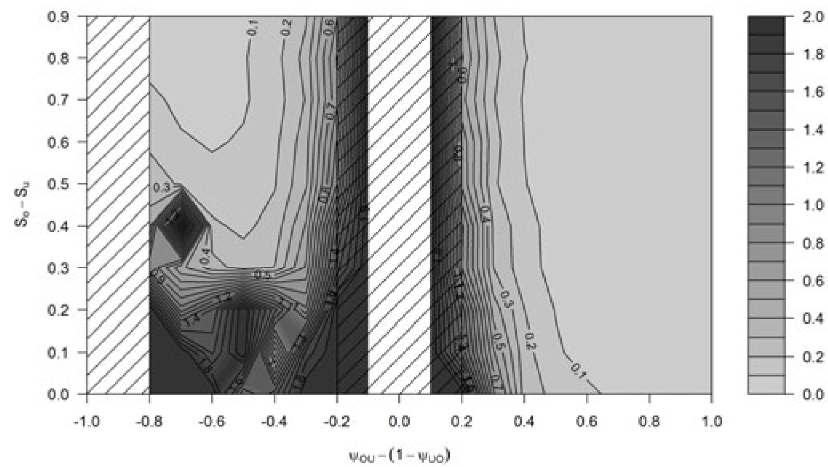
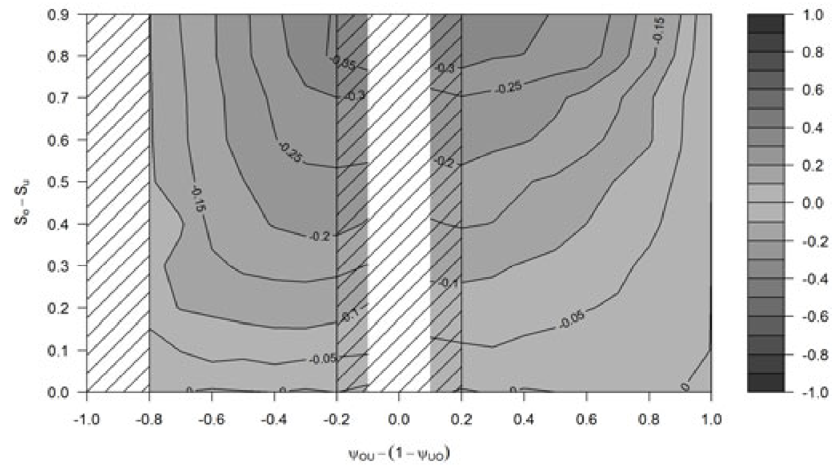


Figure 4 a) Bias and b) CV in ψ_{UO}, TE model

a)



b)

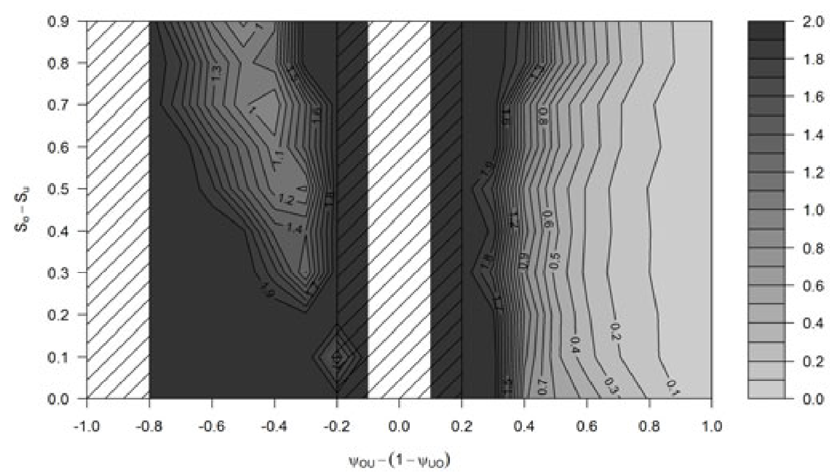
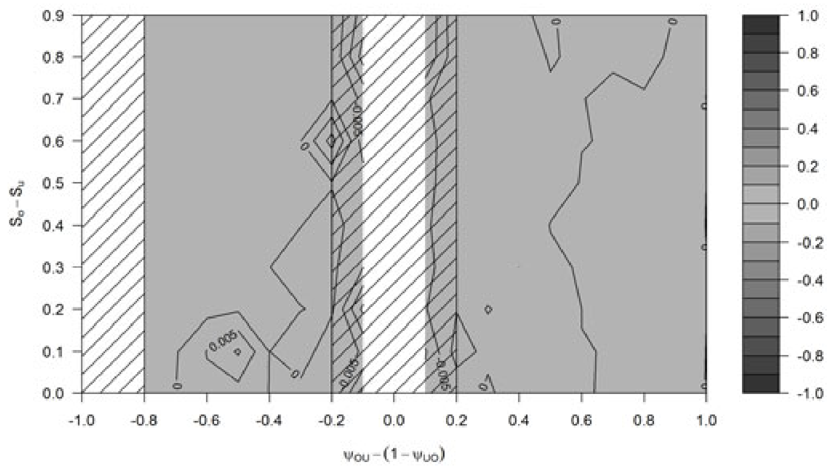


Figure 5 a) Bias and b) CV in p, TE model

a)



b)